Preemptive Scheduling of Stateful GPU Intensive HPC Applications in Kubernetes

Radostin Stoyanov - PhD Student, Scientific Computing Group

Collaboration with Adrian Reber, Senior Principal Software Engineer

Supervisor: Prof. Wes Armour

Kubernetes and HPC

Kubernetes:

- Deployment of ***loosely coupled***, containerized services
- Eventual consistency model: ***monitor & react***
- Multi-tenancy with ***shared node*** resources
- Fault-tolerance & preemption through ***termination & restart***

HPC

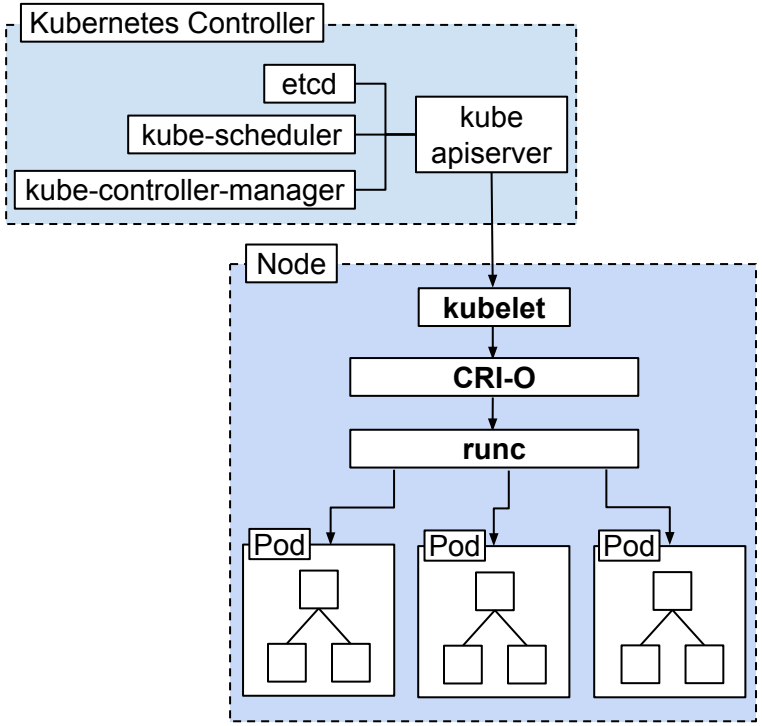
- Many applications are ***tightly coupled***
- High parallel efficiency: tasks are ***subdivided across processors***
- Job schedulers (e.g., Slurm, Torque, IBM Spectrum) allocating ***entire nodes to jobs***
- Fault-tolerance & preemption through ***checkpoint & restart***

Container Checkpointing

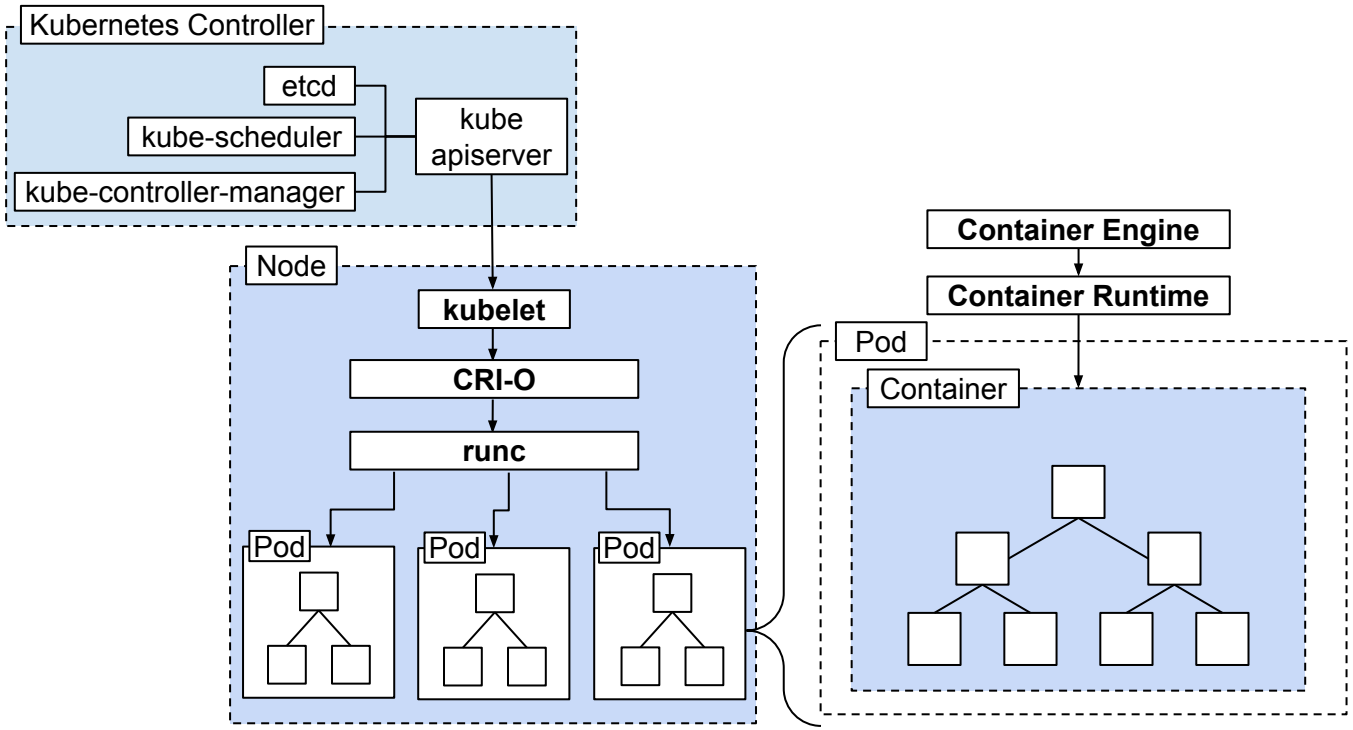
An alpha feature in Kubernetes v1.25

https://github.com/kubernetes/enhancements/issues/2008

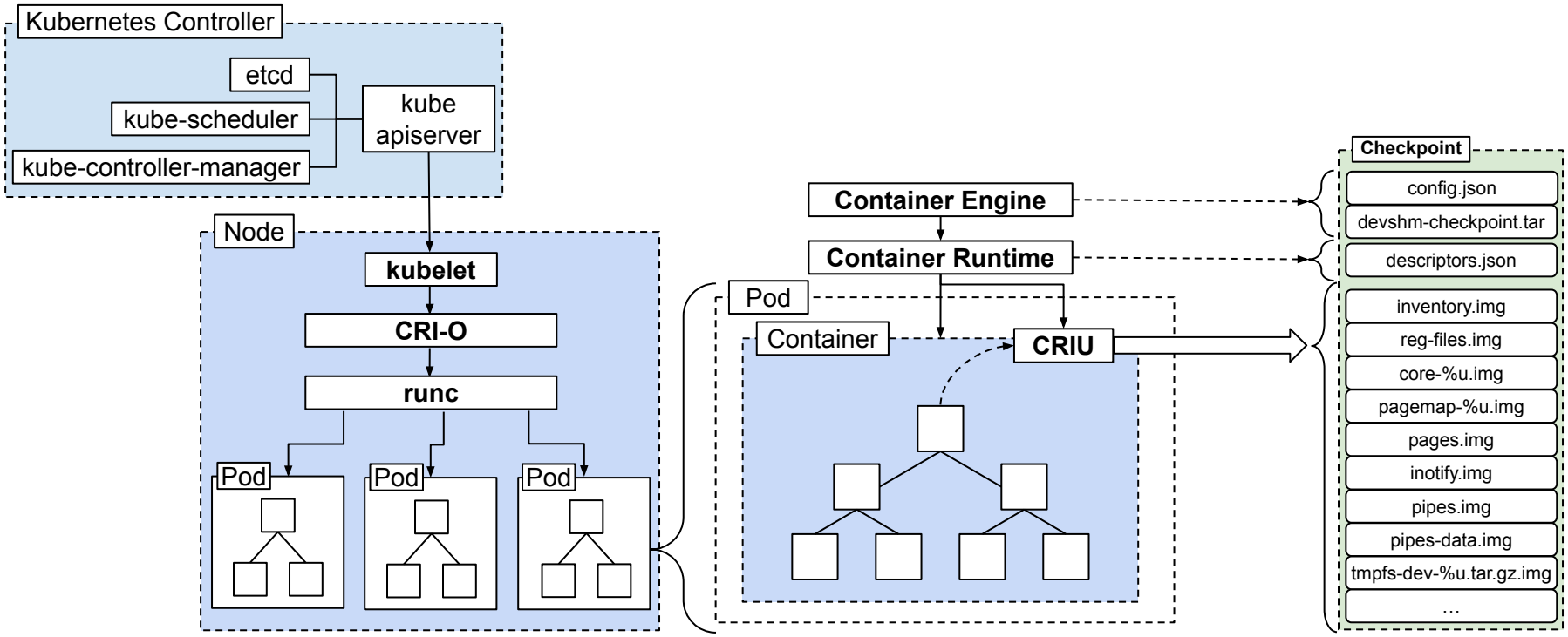
Container Checkpointing



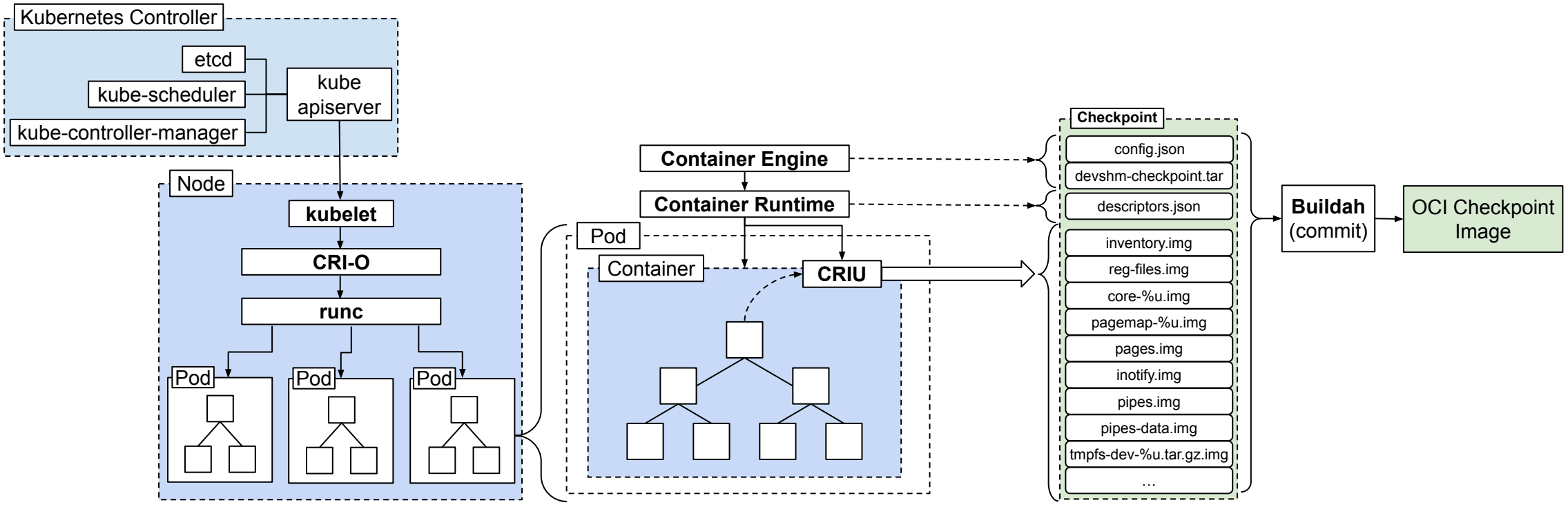
Container Checkpointing



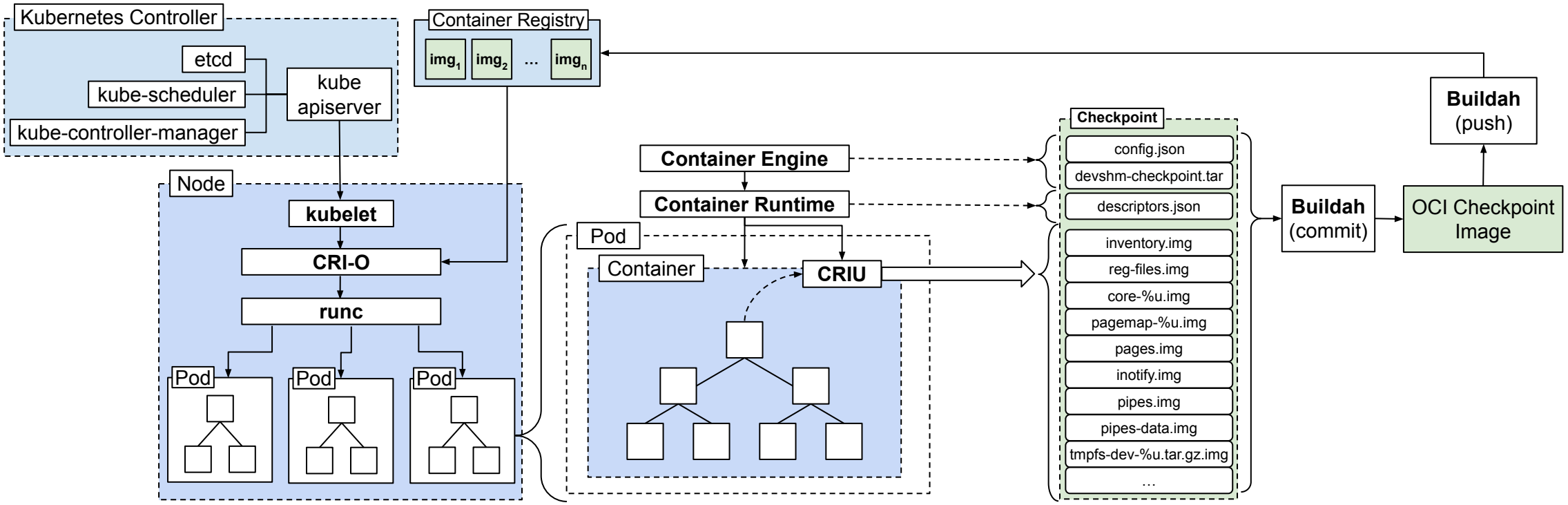
Container Checkpointing



Container Checkpointing

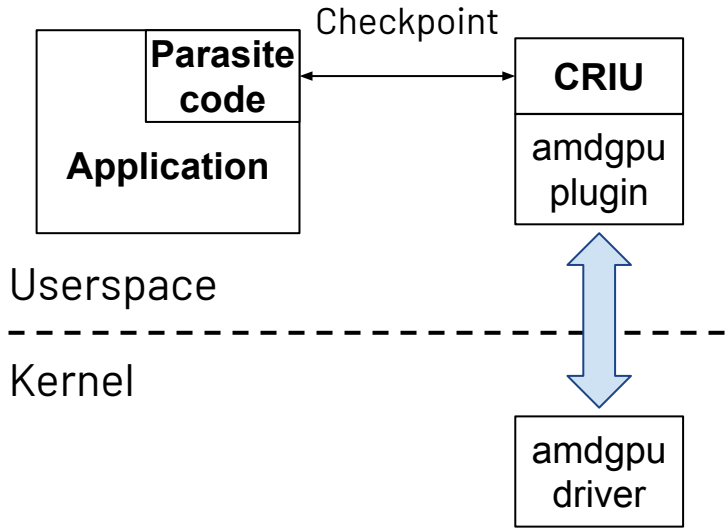


Container Checkpointing



GPU Checkpointing

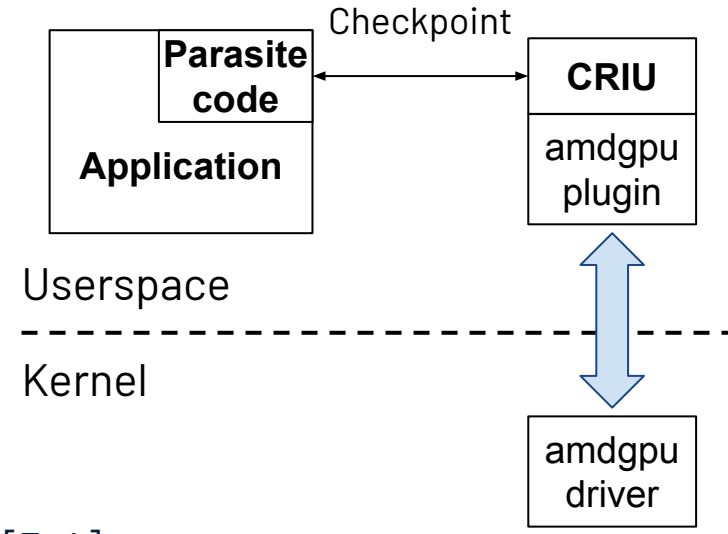
- Checkpoint support for AMD GPUs
  - AMDGPU plugin built-in as part of CRIU [1]
  - Upstream Linux kernel support via ioctl APIs [2]



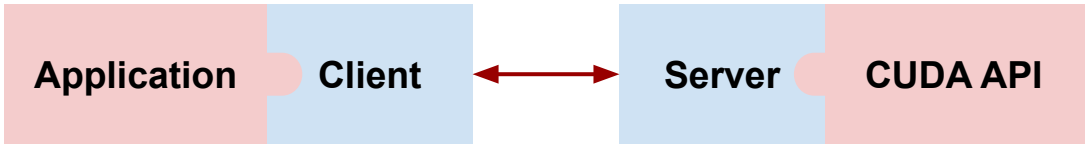
[1] **Fast Checkpoint Restore for GPUs**, Rajneesh Bhardwaj, et. al, Linux Plumbers Conference 2021
[2] **drm/amdkfd: CRIU Introduce Checkpoint-Restore APIs** (https://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git/commit/?id=36988070)
[3] **Singularity: Planet-Scale, Preemptive and Elastic Scheduling of AI Workloads**, Dharma Shukla, et. al., 2022
[4] **Cricket: A virtualization layer for distributed execution of CUDA applications with checkpoint/restart support**, Niklas Eiling, et. al., 2021

GPU Checkpointing

- Checkpoint support for AMD GPUs
  - AMDGPU plugin built-in as part of CRIU [1]
  - Upstream Linux kernel support via ioctl APIs [2]

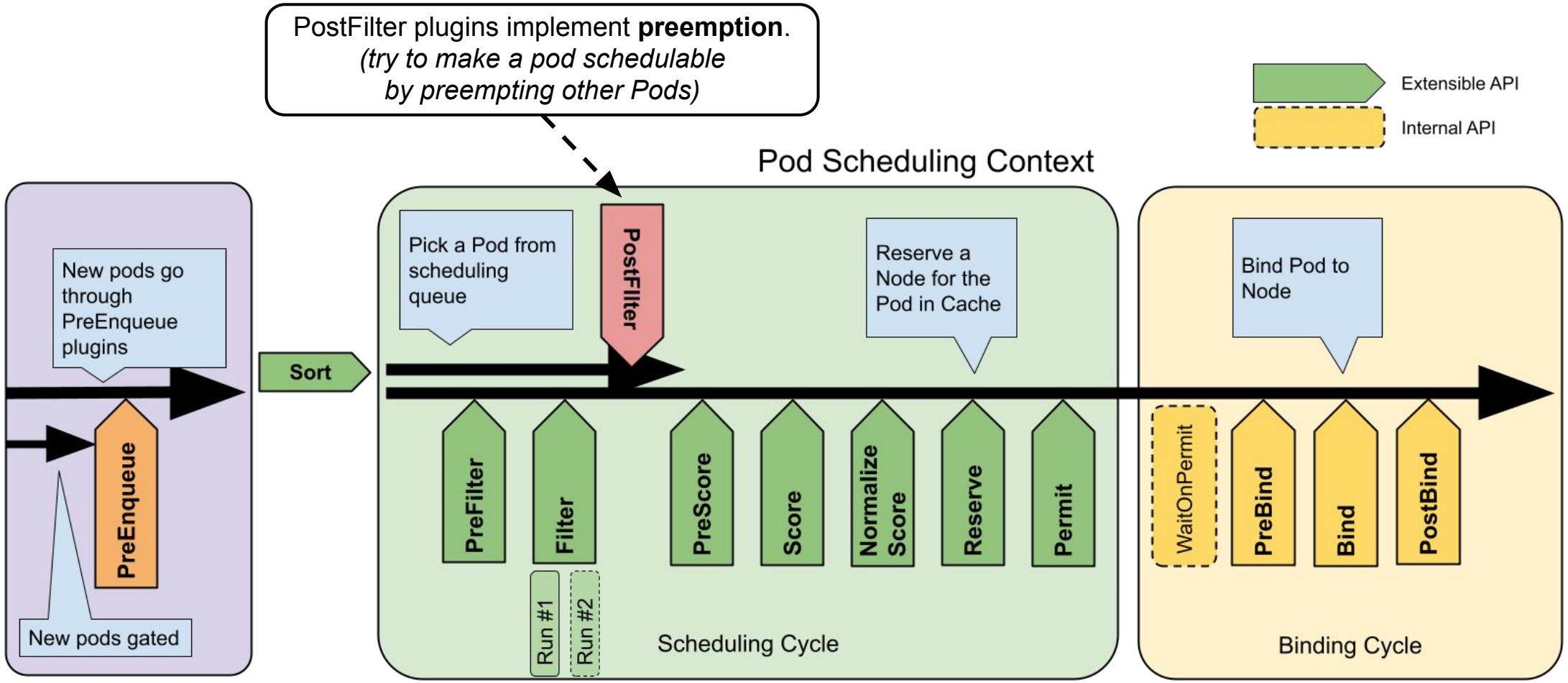


- Checkpoint support for NVIDIA GPUs with "device-proxy" [3, 4]
  - Intercepting driver API calls



[1] **Fast Checkpoint Restore for GPUs**, Rajneesh Bhardwaj, et. al, Linux Plumbers Conference 2021
[2] **drm/amdkfd: CRIU Introduce Checkpoint-Restore APIs** (https://git.kernel.org/pub/scm/linux/kernel/git/torvalds/linux.git/commit/?id=36988070)
[3] **Singularity: Planet-Scale, Preemptive and Elastic Scheduling of AI Workloads**, Dharma Shukla, et. al., 2022
[4] **Cricket: A virtualization layer for distributed execution of CUDA applications with checkpoint/restart support**, Niklas Eiling, et. al., 2021

Preemptive Scheduling



https://kubernetes.io/docs/concepts/scheduling-eviction/scheduling-framework/

Preemptive Scheduling

- Default: Pods preempt lower-priority pods (`preemptionPolicy: PreemptLowerPriority`)
- Non-preempting Pods (`preemptionPolicy: Never`)
  - will be placed in the scheduling queue ahead of lower-priority pods, but cannot preempt other pods.
- **Proposal**: Pod preemption using container checkpoint / restore

`preemptionPolicy: CheckpointLowerPriority`

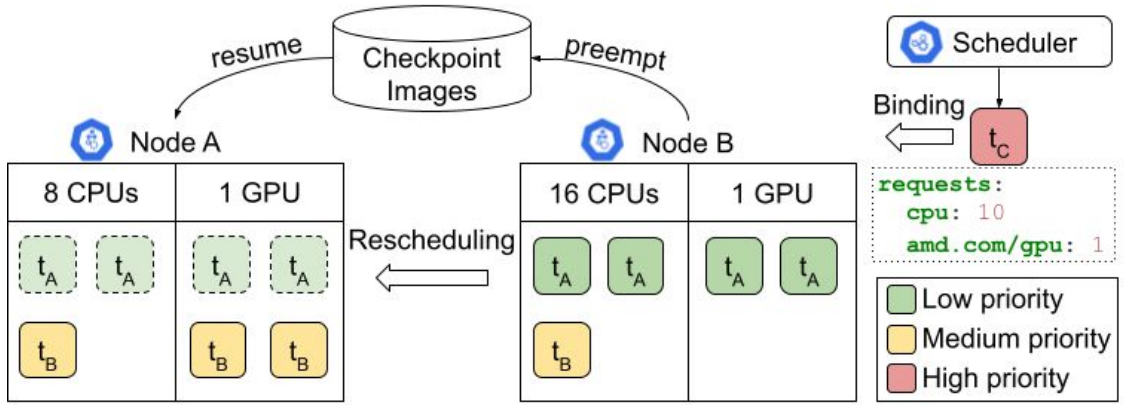
```yaml
apiVersion: scheduling.k8s.io/v1
kind: PriorityClass
metadata:
 name: checkpoint-lower-priority
value: 1000000
preemptionPolicy: CheckpointLowerPriority
globalDefault: false
description: "This priority class will preempt pods using checkpointing."
```

https://kubernetes.io/docs/concepts/scheduling-eviction/pod-priority-preemption/

Preemptive Scheduling

- Default: Pods preempt lower-priority pods (`preemptionPolicy: PreemptLowerPriority`)
- Non-preempting Pods (`preemptionPolicy: Never`)
  - will be placed in the scheduling queue ahead of lower-priority pods, but cannot preempt other pods.
- **Proposal**: Pod preemption using container checkpoint / restore

```
preemptionPolicy: CheckpointLowerPriority
```



https://kubernetes.io/docs/concepts/scheduling-eviction/pod-priority-preemption/

HPC Applications

	Size (GiB)	Memory (GiB)	GPU (GiB)	Freezing (s)	Frozen (s)	Memdump (s)	Memwrite (s)
Binomial Option Pricing	0.92	0.79	0.14	0.1	0.66	0.28	0.22
Bitonic Sort	28.4	28.29	0.14	0.1	24.9	24.5	23.9
Discrete Cosine Transform	1.3	1.09	0.14	0.1	0.51	0.35	0.30
Haar Wavelet Decomposition	0.38	0.23	0.14	0.1	0.35	0.11	0.07
Fast Walsh Transform	0.39	0.25	0.14	0.1	0.33	0.11	0.07
Floyd Warshall	5.4	5.29	0.14	0.1	1.8	1.64	1.52
Histogram	14.7	14.5	0.14	0.1	8.3	8.02	7.69
Matrix Multiplication	28.7	28.5	0.14	0.1	25.3	24.9	24.1
Recursive Gaussian	0.32	0.19	0.14	0.1	0.38	0.09	0.05
Simple Convolution	16.3	8.5	7.79	0.1	6.8	2.76	2.55

https://github.com/ROCm-Developer-Tools/HIP-Examples

Acknowledgments

- Rajneesh Bhardwaj (AMD)
- Tina Friedrich, Steven Young and Andrew Gittings (ARC Oxford)

Summary & Future work

https://criu.org/

https://github.com/checkpoint-restore/criu

https://kubernetes.io/blog/2022/12/05/forensic-container-checkpointing-alpha/